

Diversity and Complexity of Hand Movement for Autism Spectrum Disorder Intervention

Dinghuang Zhang, Carrie M Toptan, Gongyue Zhang, Shuiwen Zhao, Dalin Zhou, and Honghai Liu, *Fellow, IEEE*

School of Computing, University of Portsmouth, Portsmouth, UK

Email: {dinghuang.zhang, carrie.toptan, gongyue.zhang, shuwen.zhao, dalin.zhou, honghai.liu}@port.ac.uk

Abstract—Flexibility and adaptability described in individuals with Autism Spectrum Disorders (ASD) refer to Stereotypical Motor Movements (SMM) and social interaction deficits, both of which are important symptoms of ASD. Inspired by the most recent psychological research by Jonge-Hoekstra, this paper aims to distinguish hand movement with two quantitative metrics extracted by the mid-layer of a supervised convolutional gesture recognition network, named diversity and complexity. Particularly, diversity indicates adaptability and complexity indicates flexibility. The utilisation of both metrics shows great potential for hand movement analysis with a particular emphasis on ASD intervention.

Keywords—Autism, ASD, flexibility, adaptability, skeleton-based gesture recognition, CNN, recurrence analysis.

I. INTRODUCTION

Autism Spectrum Disorders (ASD) is an early onset neurodevelopmental disorder, characterised by deficits in social interaction and communication and repetitive, restricted behaviours or interests. Although there is no cure for ASD, research shows that early screening and diagnosis can help increase a child's chances of benefiting from early intervention [1].

The difficulty with the current clinical diagnosis, apart from the lack of trained clinicians, is that the traditional scale based behaviour observation is both time-consuming and insufficient, as children are extremely different from each other and the behaviours are seen in children often depend on many non-autistic related factors, including cognitive functioning and age, so some of the traits that human intuition and statistical analysis expect to uniformly distinguish autistic from non-autistic may not exist. Particularly in the evaluation of physical behaviour in children with autism, therefore the deficient is not only in standardising the assessment, But also difficult in determining subtle heterogeneous behaviour. To compensate for this possible deficient of scale-based methods and provide a general tool for movement analysis, quantitative analysis of movement based on deep learning algorithms and computer vision can significantly improve the recognition accuracy while standardising the quantification process. And this movement quantification technique has great potential to be used as an important part of technology-assist ASD diagnosis, due to its high repeatability, high resolution, and high sensitivity.

Hand movement is strongly associated with the determination of non-verbal expressive and repetitive stereotyped behaviours in the diagnosis of ASD. A very recent [2]Acta Psy-

chologica research by Jonge-Hoekstra propose two variability measures for children hand movement in spatial-temporal task properties: Complexity indicates flexibility, Diversity indicates adaptation. Their work indicates the possibility of mathematically quantify hand movement. Furthermore, according to the medical researches [3] [4] [5] [6] the flexibility and the adaptation behaviours are both strong predictor in ASD. Where the flexibility in hand behaviours in individuals with ASD refers to the Restricted Repetitive Behaviours (RRB) and Stereotype Behaviours (SB) [7] [8] [9]. And adaptive functioning encompasses those behaviours critical to living independently, including daily living skills, social skills, and communication skills. Furthermore, children with Autism Spectrum Disorders (ASD) display adaptive behaviours impairments across all domains [10]. Furthermore, Motor difficulties is currently not included in the diagnostic criteria for ASD. However the proportion of children with ASD having motor difficulties is 86.9% [11] [12]. Although motor difficulties are not included in the diagnostic criteria for ASD, it is included in the ICF criteria for ASD which is a tool designed to be a standardised and internationally accepted tool for ASD functioning [13].

There have been several approaches that have attempted to apply deep learning techniques to detect and diagnose ASD [14] [15] [16] [17]. Many recent studies have used machine learning classifiers to automatically detect stereotypical motor movements (SMM) in individuals with ASD [18] [19] [20] [21]. But the machine learning classifiers are establishing an end-to-end connection between behaviour and certain types of SMM. This end-to-end connection still based on manual behavioural observation, for instance, the dataset is still labelled by the handcraft rules. And this qualitative classification result can not describe the behaviour quantitatively. Indifference to these methods, Ulf [22] utilises recurrence quantification analyses (RQA) to capture non-linear dynamics of SMM through wearable accelerators. Their study also supports our attempt in establishing movement categories through similar skeleton features.

The hand movement quantitative measurement in Jonge-Hoekstra's approach is based on hand-crafted categories of hand movement, then they apply Shannon entropy and RQA to estimate the diversity and complexity based on the total of 14 different categories. Different from their hand-crafted categorisation we attempt to use deep learning algorithms to find the universal sub-features that appear in hand movements

that constitute the hand behaviour and use these sub-features as a basis for movement categorisation. Since the sub-features that constitute movement are universal, the gestures of autism are also composed of universal sub-features. Theoretically, both normal and abnormal hand movements are made up of a series of joint variations. So, that kind of hand motion sub-features can learn from any skeleton-based hand motion dataset and use as a prior-knowledge for ASD hand behaviour analysis.

The skeleton-based Action recognition, with the convenience of obtaining skeleton information, has become a hot topic [23] [24] [25] [26]. Similar to other skeleton-based gesture recognition approaches [27] [28] [29], our proposed method also use a skeleton data enforce module to enrich the spatial-temporal information. The strategy to explore the sub-features by applying gesture recognition on the target skeleton dataset, compression of the intermediate layer's output feature map using an hourglass structure convolutional network, then mapping this feature map to categorized sub-features. Cause the gesture recognition is built on this feature map the high accuracy of the gesture recognition the higher credibility about the feature map and corresponding sub-features. In the training process, the model is going to define a series of the convolutional layer to explore the universal patterns which composite movement.

To sum up, to achieve measurement of complexity and diversity, we want to utilise deep learning algorithms to explore the universal joint sub-features from a hand movement dataset and implement classification based on these sub-features for measurement computation. In this paper, we propose a CNN based motion feature explore model for skeleton-based hand gesture recognition, where the sub-feature classification is based on co-occurrence feature exploration layer's output will explain in section III. Considering the characteristics of convolution and hand movement, a hand skeleton data reorganisation module is designed to compensate for spatial-temporal related information to enhance the robustness of the model. The main contributions of this work are summarised as follows:

- We define the two measurements of hand motion diversity and complexity in terms of skeleton data.
- We propose a deep learning hand movement measurement model utilising the sub-features to estimate diversity and complexity.
- The proposed skeleton-based behaviour quantification model has great potential in hand behaviour analysis, it paves the way for future use in ASD intervention.

The rest part of this paper is organised as: Section II introduce the proposed data reorganising strategy, gesture recognition, sub-feature-based classification methods and quantification method; Section III, dataset, evaluation and related experiments setting details; Section IV conclusion.

II. METHODOLOGY

A. Data Pre-Processing



Fig. 1. (a) Hand joint model. (b) Graph-based hand skeleton organisation. FJ: first joint. SJ: Second joint. Note there will be no angle value at wrist and tip rows.

The graph-based skeleton data for each joint is a 3-dimensional vector (x, y, z) . According to the finger structure, the connected joints are vectors, for each joint, the angle R is the adjacent two vectors, is:

$$R_i^t = \cos^{-1} \left(\frac{J_i^t \vec{J}_{i+1}^t \times J_{i-1}^t \vec{J}_i^t}{|J_i^t \vec{J}_{i+1}^t| \times |J_{i-1}^t \vec{J}_i^t|} \right) \quad (1)$$

Where, t : time-series; for each $J \in (x, y, z)$, $i, i+1$ and $i-1$ is the adjacent joints in columns in Fig. 1(b). The skeleton data for each frame for each joint after data reorganisation contains 4-dimension information (x, y, z, R) and format as Fig. 1(b).

B. Temporal Difference of Skeleton Movement

Temporal difference for time t is defined, m is:

$$m^t = s^{t+1} - s^t \quad (2)$$

where, s^t is the skeleton matrix in time: t ;

$$M = (m^1, m^2, \dots, m^{N-1}) \quad (3)$$

This spatial-temporal change represents method Eq.3 ignores the specific positions of the joints and only focuses on the temporal difference (movement) of the joints, which enhance the spatial-temporal ability of the CNN network.

C. Sub-features Estimation

To generalise the model and improve portability, we want the model can explore the 'phonemes' of hand movement. And the recognise/classification is based on the explored 'phonemes' (general characteristics of hand movement). Furthermore, if the model with very robust outcome in testing, we can treat the 'phonemes' output as a different class of hand movement patterns, that how the sub-feature classification built. The difficulty lies in the choice of network structure and how to prove the reliability of sub-feature classification. In the input of skeleton sequence using image-based representation, CNN does have a great advantage to recognise the special

structure in the graphical sequence, but the challenge is how to effectively capture the spatial-temporal information. For transfer learning, the network must obtain enough gesture recognition accuracy on the test set to show that the sub-feature classification layer is sufficiently reliable.

Based on the hand motion array M obtained by skeleton data pre-Processing, the network is expecting to learn the commonalities from the connected joints of the same finger and the adjacent joints of different fingers in this skeleton sequence. We design two different shapes of convolution kernel to extract the possible features of the same joints of each finger and different fingers. To make the feature map that can contain as much information as possible, we are very restrained in adding a pooling layer in co-occurrence feature explore layers.

D. Sub-feature Categorisation

Through gesture recognition training process, the categorisation sub-feature will utilise the network's feature explore lays and corresponding weights, to calculate the feature map. For this layer's output, the feature map is a two-dimensional matrix N with size (t, n) , where t refers to the number of skeleton sequence and n is the dimension of the sub-feature. For each time t a threshold th will apply binarization of the data:

$$Binary = \begin{cases} 1 & \text{if } N[i] > t \\ 0 & \text{if } N[i] < t \end{cases} \quad (4)$$

After the two-dimensional matrix, N has applied binarization, for each time t the binary number will be converted to a decimal number to indicate the category of movement, thus the categories of an action sequence on the time series are established, so that the diversity and complexity can be measured. In our experiment, the feature map N size is set to 6, which means after categorisation we expect 64 categories sub-feature in total.

E. Jonge-Hoekstra based Measurement

Diversity means adaption, which indicates the functional adjustment to new task demands, based on the frequencies of the different behavioural categories of hand movements duration. Their measure Diversity by calculating Shannon entropy on the frequency distribution of the duration and occurrence of behavioural categories of the time-series. The diversity value only related to the amount of variability of the participants' gestures.

Complexity means flexibility, which indicates functional flexibility when changes in task demands. It is a measurement by performing RQA on the time series of gestures, for instance, RQA is a nonlinear time series analyse method based on the notion of recurrence. These recurrences are represented in a Recurrence Plot (RP), for categorical time series (the block structures in the RP), which reflects the dynamic, temporal organisation of a system. The complexity is also measured by Shannon entropy, based on the block structures in the RP. Where the Shannon entropy of frequency distribution of

the block structures in the RP reflects the complexity of the deterministic structure in the system [30].

III. EXPERIMENT AND RESULTS

A. Hand Movement Dataset

TABLE I
DHG14/28 DATASET

Dynamic Hand Gesture	
Gestures	14 Class(single finger /whole hand)
Joint Number	22 Joints
Sequence	2800
Data type	RGB, Depth
Label	Skeleton(2D, 3D), Gesture
Participants	20
Handedness	Right-handed

The Dynamic Hand gesture 14/28 (DHG-14/28) [31] is a public dynamic hand gesture dataset, which sequences of 14 hand gestures performed 5 times by 20 participants in 2 finger configurations: using one single finger or the whole hand. Each gesture is performed 5 times (Essai) by 20 participants in 2 way, resulting in 2800 sequences TABLE I. The coordinates of 22 hand joints in the 3D world space are provided per frame, forming a full hand skeleton.

All of these 14 finger movements/gestures were involved in this study, but our movement quantification method is not limited to those 14 gestures. That because our proposed deep learning model utilizing those 2800 skeleton sequence to explore the universal patterns for hand movement to estimate the sub-feature. And obviously, the movement patterns extracted from the skeletal movements of these 14 gestures have certain limitations, but at this stage, we think these patterns can meet the requirement.

B. Implementation Details

1) *Network structure*: Our network structures are illustrated in Fig. 2 For the batch training set, the input data in the time series for each gesture are resized to (batch-size, 4, 5, 5*t). Our proposed deep learning network apply a convolutional layer with stride 5, this design is based on the input shape, as we are not performing the skeleton sequence in time series to channels, but we expand the last dimension of the input matrix. We use this method to realize the processing of indefinite length sequences through convolutional layers without using LSTM, through utilizing these special convolution strategies, realize the compression of each frame of skeleton data in chronological order. Such a strategy is designed to enable the network during the gesture recognition training process make the learned convolution kernel in the co-occurrence feature exploration part more targeted at the general movement features (sub-feature).

2) *Network Training*: The proposed network is implemented based on the PyTorch platform. The Adam [32] optimizer with a scheduled learning rate with initial set to 0.001 with step-size 25, gamma 0.5, employed in the training process.

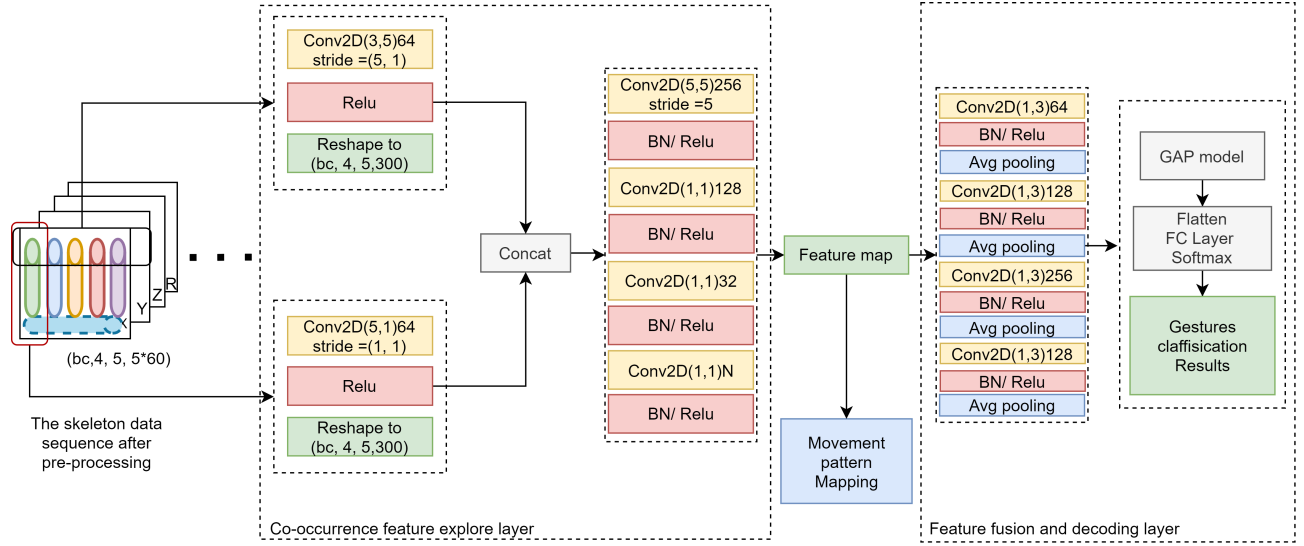


Fig. 2. Network structure illustration. The network can be divided into 2 key parts. 1) co-occurrence feature explores. 2) Feature fusion and decode.

C. Performance on DHG14 gesture classification

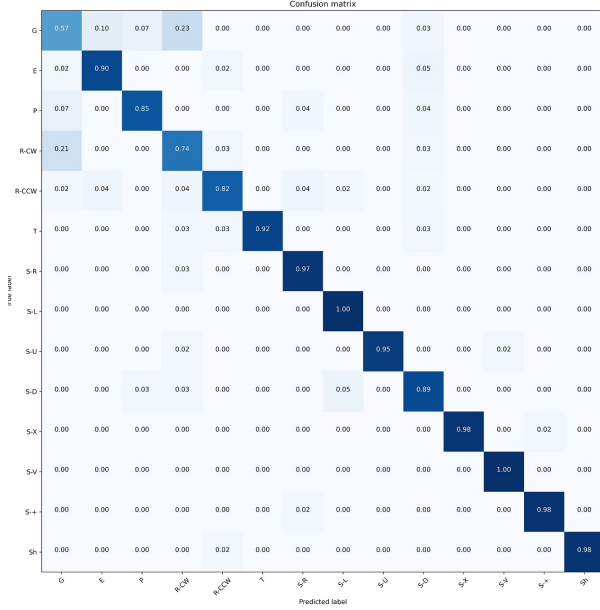


Fig. 3. The confusion matrix of the proposed approach for DHG-14

The action recognition results of the DHG-14 dataset are presented in Table II and to better understand the behaviour of our approach according to the recognition per class more details are listed in the confusion matrix Fig. 3.

We compare our method with the state-of-the-art methods on the DHG Dataset in TABLE.II. Our method achieves 90.5% accuracy. Our result illustrates the superiority of some previous methods like SoCJ+HoHD+HoWR, CNN+LSTM and STA-Res-TCN on the DHG dataset. Compare with the most advanced methods like ST-GCN and DG-STA, although our method did not beyond their recognition accuracy, but our

gesture recognition results have provided sufficient credibility for the sub-features.

TABLE II
COMPARE WITH STATE OF ART

Method	DHG 14 Result
SoCJ+HoHD+HoWR [33]	83.1
CNN+LSTM [34]	85.6
Res-TCN [35]	86.9
STA-Res-TCN [35]	89.2
ST-GCN [36]	91.2
DG-STA [37]	91.9
Ours Method	90.5

D. RQA Calculation

We utilise the pyRQA [38] tool to calculate the entropy.

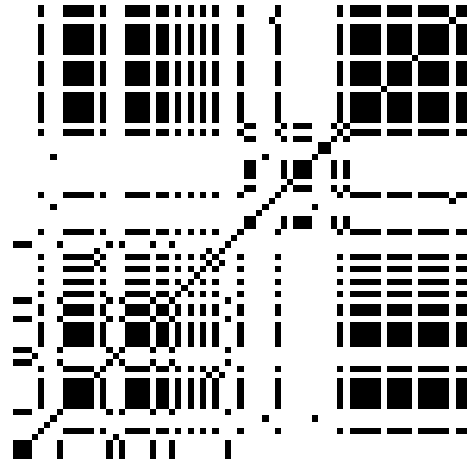


Fig. 4. Recurrent plot for object 1 Gesture grab in DHG-14 dataset perfume by pyRQA with following settings

In RQA analysis we set dimension parameter to 1, time delay τ to 1 and set similarity measure to Euclidean Metric and

neighbourhood parameter with fixed radius with 1 in pyRQA settings.

E. Diversity and Complexity Result and analysis

We apply the proposed hand measurement method on DHG dataset, through comparing different class of gestures, object, and finger number, the result as follow.

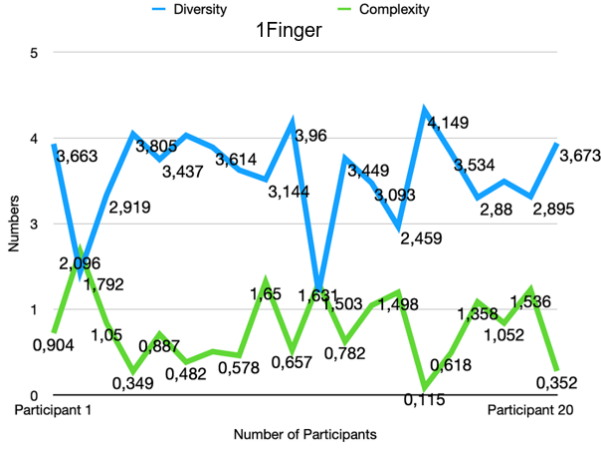


Fig. 5. Gesture[1]:Grab, 1 Finger, Essai:1

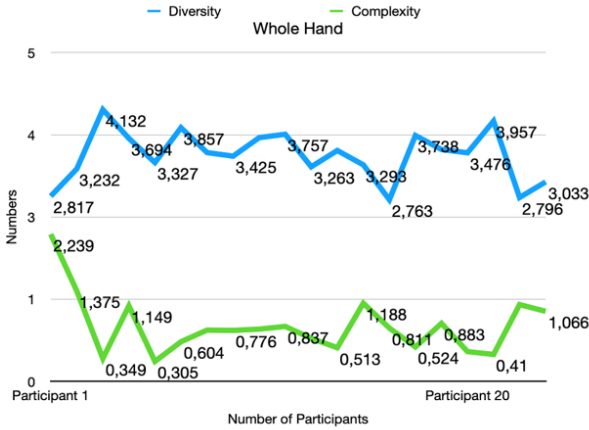


Fig. 6. Gesture[1]:Garb, Whole Hand, Essai:1

Fig. 5 shows that Diversity and Complexity are very close to each other at two points (Diversity: 2.096, Complexity: 1.792; Diversity: 1.631, Complexity: 1.503) when gesturing using one finger. This Figure also shows the discrepancies between Diversity and Complexity on the other points. In comparison Fig. 6 shows the discrepancies between Diversity and Complexity when using the whole hand. These two Figures (Fig. 5 6) show how the results differ in Diversity and Complexity in hand movements when using one finger compared to using the whole hand.

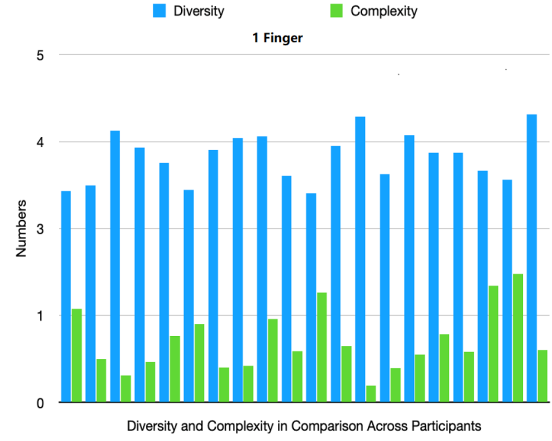


Fig. 7. Gesture[8]:Swipe Left, 1 Finger, Essai:1

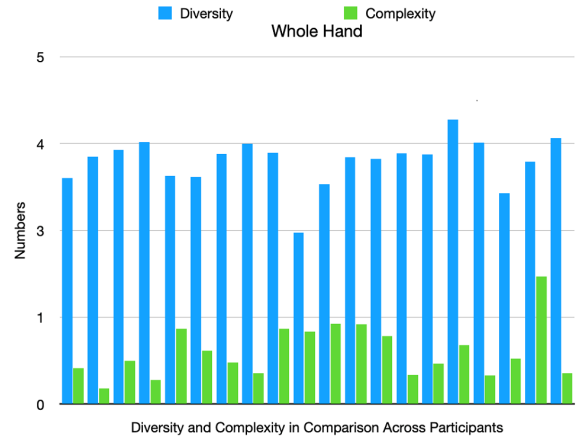


Fig. 8. Gesture[8]:Swipe Left, Whole Hand, Essai:1

In Fig. 7 it is clearly shown that Complexity and Diversity differ enormously in all participants. Furthermore, the Figure shows that Diversity is very high compared to Complexity when using one finger. Additionally, Fig. 8 shows the huge discrepancies between Diversity and Complexity when using the whole hand, which is very similar when using only one finger.

We also find a strong, negative, linear relationship for movements perform with one finger when plotting Diversity and Complexity as x and y-axis in Cartesian coordinate, but the Whole Hand shows a non-linear, negative relationship.

We conclude that with the same gesture use the same finger when comparing different participants, the diversity change is small, but the complexity is different. Comparing a single finger with the whole hand when performing the same action with the same participant, the diversity and complexity are increased. However, these characteristics are not absolute. Adjusting both the RQA setting and categorise strategy will cause great changes in the result, more experiments are needed to fine-tune the Network and the RQA parameters. Further research could investigate the network on more datasets, apply the network to real-time use to offer more use-age scenarios. And work with professionals with a psychology background

apply this indicator to the diagnosis and treatment of autism under professional guidance.

IV. CONCLUSION AND FUTURE WORK

In this paper, a skeleton-based quantification method was proposed to distinguish hand movement. An Hourglass network has been enforced with hand skeleton data and outperformed the classic gesture recognition model. Particularly, two quantitative metrics, named diversity and complexity, were defined through categorisation of movement by sub-features to describe hand movement. The proposed method provides a general framework for tasks involving hand movement quantification. However, this work is still in a preliminary stage, its potential in ASD intervention and clinical rehabilitation will be examined in future work. Further experiments are included to prove the feasibility of diversity and complexity across scenarios and how to minimize the influence due to individual difference.

REFERENCES

- [1] B. Reichow, E. E. Barton, B. A. Boyd, and K. Hume, "Early intensive behavioral intervention (eibi) for young children with autism spectrum disorders (asd)," *Cochrane Database Syst. Rev.*, no. 10, 2012.
- [2] L. de Jonge-Hoekstra, S. van der Steen, and R. Cox, "Movers and shakers of cognition: Hand movements, speech, task properties, and variability," *Acta Psychol.*, vol. 211, 10 2020.
- [3] K. Y. K. Y. K. S. M. C. M.-N. C. P. V. P. C. W. C. Y. M. F. E. Elsabbagh M, Divan G, "Global prevalence of autism and other pervasive developmental disorders," *Autism Res*, pp. 5(3):160–79, 2012.
- [4] L. Kenworthy, L. Case, M. Harms, A. Martin, and G. Wallace, "Adaptive behavior ratings correlate with symptomatology and iq among individuals with high-functioning autism spectrum disorders," *J. Autism Dev. Disord.*, vol. 40, pp. 416–23, 04 2009.
- [5] M. Lewis and S.-J. Kim, "The pathophysiology of repetitive behavior," *J. Neurodev. Disord.*, vol. 1, pp. 114–32, 06 2009.
- [6] L. Bishop, J. Hong, L. Smith, R. Makuch, J. Greenberg, and M. Mailick, "Characterizing objective quality of life and normative outcomes in adults with autism spectrum disorder: An exploratory latent class analysis," *J. Autism Dev. Disord.*, vol. 46, 08 2016.
- [7] K. AV, "Parent expectations mediate outcomes for young adults with autism spectrum disorder," *J. Autism Dev. Disord.*, vol. 46, 05 2016.
- [8] C. Joyce, E. Honey, S. R. Leekam, S. L. Barrett, and J. Rodgers, "Anxiety, intolerance of uncertainty and restricted and repetitive behaviour: Insights directly from young people with asd," *J. Autism Dev. Disord.*, vol. 47, no. 12, pp. 3789–3802, 2017.
- [9] D. A. Baribeau, S. Vigod, E. Pullenayegum, C. M. Kerns, P. Mirenda, I. M. Smith, T. Vaillancourt, J. Volden, C. Waddell, L. Zwaigenbaum, et al., "Repetitive behavior severity as an early indicator of risk for elevated anxiety symptoms in autism spectrum disorder," *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 59, no. 7, pp. 890–899, 2020.
- [10] R. Militeri, C. Bravaccio, C. Falco, C. Fico, and M. t. Palermo, "Repetitive behavior in autistic disorder," *Eur. Child Adolesc. Psychiatry*, vol. 11, pp. 210–8, 11 2002.
- [11] M. K. Licari, G. A. Alvares, K. Varcin, K. L. Evans, D. Cleary, S. L. Reid, E. J. Glasson, K. Bebbington, J. E. Reynolds, J. Wray, et al., "Prevalence of motor difficulties in autism spectrum disorder: Analysis of a population-based cohort," *Autism Res.*, vol. 13, no. 2, pp. 298–306, 2020.
- [12] A. N. Bhat, "Is motor impairment in autism spectrum disorder distinct from developmental coordination disorder? a report from the spark study," *Physical therapy*, vol. 100, no. 4, pp. 633–644, 2020.
- [13] S. Bölte, S. Mahdi, P. J. de Vries, M. Granlund, J. E. Robison, C. Shulman, S. Swedo, B. Tonge, V. Wong, L. Zwaigenbaum, et al., "The gestalt of functioning in autism spectrum disorder: Results of the international conference to develop final consensus international classification of functioning, disability and health core sets," *Autism*, vol. 23, no. 2, pp. 449–467, 2019.
- [14] H. Cai, Y. Fang, Z. Ju, C. Costescu, D. David, E. Billing, T. Ziemke, S. Thill, T. Belpaeme, B. Vanderborght, et al., "Sensing-enhanced therapy system for assessing children with autism spectrum disorders: A feasibility study," *IEEE Sens. J.*, vol. 19, no. 4, pp. 1508–1518, 2018.
- [15] Z. Wang, J. Liu, K. He, Q. Xu, X. Xu, and H. Liu, "Screening early children with autism spectrum disorder via response-to-name protocol," *IEEE Trans Industr Inform.*, vol. 17, no. 1, pp. 587–595, 2019.
- [16] H.-L. Cao, P. G. Esteban, M. Bartlett, P. Baxter, T. Belpaeme, E. Billing, H. Cai, M. Coeckelbergh, C. Costescu, D. David, et al., "Robot-enhanced therapy: Development and validation of supervised autonomous robotic system for autism spectrum disorders therapy," *IEEE Robot Autom Mag*, vol. 26, no. 2, pp. 49–58, 2019.
- [17] J. Li, Y. Zhong, J. Han, G. Ouyang, X. Li, and H. Liu, "Classifying asd children with lstm based on raw videos," *Neurocomputing*, vol. 390, pp. 226–238, 2020.
- [18] N. Gonçalves, S. Costa, J. Rodrigues, and F. Soares, "Detection of stereotyped hand flapping movements in autistic children using the kinect sensor: A case study," in *ICARSC*, pp. 212–216, 2014.
- [19] U. Großekathöfer, N. V. Manyakov, V. Mihajlović, G. Pandina, A. Skalkin, S. Ness, A. Bangerter, and M. S. Goodwin, "Automated detection of stereotypical motor movements in autism spectrum disorder using recurrence quantification analysis," *Front. Neuroinform.*, vol. 11, p. 9, 2017.
- [20] N. M. Rad and C. Furlanello, "Applying deep learning to stereotypical motor movement detection in autism spectrum disorders," in *ICDMW*, pp. 1235–1242, IEEE, 2016.
- [21] L. Sadouk, T. Gadi, and E. H. Essoufi, "A novel deep learning approach for recognizing stereotypical motor movements within and across subjects on the autism spectrum disorder," *Comput. Intell. Neurosci.*, vol. 2018, 2018.
- [22] U. Großekathöfer, N. Manyakov, V. Mihajlović, G. Pandina, A. Skalkin, S. Ness, A. Bangerter, and M. Goodwin, "Automated detection of stereotypical motor movements in autism spectrum disorder using recurrence quantification analysis," *Front. Neuroinform.*, vol. 11, 02 2017.
- [23] S. Zhang, Y. Yang, J. Xiao, X. Liu, Y. Yang, D. Xie, and Y. Zhuang, "Fusing geometric features for skeleton-based action recognition using multilayer lstm networks," *IEEE Trans. Multimed.*, vol. 20, no. 9, pp. 2330–2343, 2018.
- [24] V. Choutas, P. Weinzaepfel, J. Revaud, and C. Schmid, "Potion: Pose motion representation for action recognition," in *CVPR*, pp. 7024–7033, 2018.
- [25] M. Liu, H. Liu, and C. Chen, "Enhanced skeleton visualization for view invariant human action recognition," *Pattern Recognit*, vol. 68, 03 2017.
- [26] F. Yang, S. Sakti, Y. Wu, and S. Nakamura, "Make skeleton-based action recognition model smaller, faster and better," 2020.
- [27] W. Nie, W. Wang, and X. Huang, "Srnet: Structured relevance feature learning network from skeleton data for human action recognition," *IEEE Access*, vol. 7, pp. 132161–132172, 2019.
- [28] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," 2018.
- [29] W. Zheng, L. Li, Z. Zhang, Y. Huang, and L. Wang, "Relational network for skeleton-based action recognition," 2019.
- [30] G. Leonardi, "A method for the computation of entropy in the recurrence quantification analysis of categorical time series," *Physica A*, vol. 512, pp. 824–836, 2018.
- [31] Q. De Smedt, H. Wannous, and J. Vandeborbe, "Skeleton-based dynamic hand gesture recognition," in *CVPRW 2016*, pp. 1206–1214, 2016.
- [32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.
- [33] Q. De Smedt, H. Wannous, and J.-P. Vandeborbe, "Skeleton-based dynamic hand gesture recognition," in *CVPRW*, pp. 1–9, 2016.
- [34] J. C. Nunez, R. Cabido, J. J. Pantrigo, A. S. Montemayor, and J. F. Velez, "Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition," *Pattern Recognit*, vol. 76, pp. 80–94, 2018.
- [35] J. Hou, G. Wang, X. Chen, J.-H. Xue, R. Zhu, and H. Yang, "Spatial-temporal attention res-tcn for skeleton-based dynamic hand gesture recognition," in *ECCV*, pp. 0–0, 2018.
- [36] Y. Li, Z. He, X. Ye, Z. He, and K. Han, "Spatial temporal graph convolutional networks for skeleton-based dynamic hand gesture recognition," *EURASIP J Image Video Process*, vol. 2019, no. 1, pp. 1–7, 2019.
- [37] Y. Chen, L. Zhao, X. Peng, J. Yuan, and D. N. Metaxas, "Construct dynamic graphs for hand gesture recognition via spatial-temporal attention," *arXiv preprint arXiv:1907.08871*, 2019.